

NUEVA TÉCNICA DE CORRESPONDENCIA DE IMÁGENES PARA LA RECONSTRUCCIÓN DE ATAURIQUES

A new CBIR technology to help reassembling moorish ornamental carvings
(atauriques)

FRANCISCO JAVIER MOLINA*, JAVIER M. MORA*, JULIO BARBANCHO*
y CARLOS LEÓN DE MORA*

RESUMEN Este trabajo presenta una nueva tecnología de correspondencia de imágenes especialmente concebida para facilitar la reconstrucción de atauriques a partir las imágenes de sus fragmentos. El método localiza piezas similares en base a su contenido ornamental, ignorando la forma de su fractura. Para ello se propone una técnica de texturización que realza este contenido. Además, se propone una modificación del algoritmo de cálculo de características de Fourier clásico para mejorar la tasa de éxito y preservar la información de escala. Los resultados demuestran que esta tecnología es adecuada para filtrar o reducir el número de candidatos en un proceso de restauración basado en técnicas de reconstrucción por puzles.

Palabras clave: CBIR, Reconstrucción, Texturas.

ABSTRACT In this paper, we present a new Content Based Image Retrieval technology specially designed to help reassemble archaeological pieces. The method finds similar fragments based on the fragment content not the shape of the fracture. A texturization method is proposed to enhance this content. Furthermore, classic Fourier features computation is modified to increase the success matching, and preserve the scaling information. The results show that this technology is suitable to filter candidates in a puzzle-solving tool when the number of fragments is huge.

Key words: CBIR, Reassembling, Puzzle solving, Textures.

* Departamento de Tecnología Electrónica. Universidad de Sevilla. *fjmolina@us.es, jmmora@us.es, jbarbancho@us.es, cleon@us.es*

Fecha de recepción: 14-10-10. Fecha de aceptación: 20-01-11

1. INTRODUCCIÓN

Reensamblar fragmentos para recomponer imágenes y objetos es un problema muy común en arqueología (Carlini *et al.*, 2008). En la actualidad existen muchos algoritmos que, a partir de las fotografías de dichos fragmentos, resuelven el problema de la reconstrucción como si se tratase de un puzzle. Para ello, van comparando por pares todas las piezas siguiendo un proceso iterativo con el que se recorren todos los casos de forma exhaustiva. La comparación suele tener en cuenta propiedades del fragmento como la forma de la fractura, su contenido (color, dibujos, figuras, etc.), o ambos. En el caso de objetos como vasijas o estatuas, los modelos 3D ayudan a reducir el número de comparaciones y a mejorar los resultados de la reconstrucción.

La mayoría de los algoritmos basados en estrategias de resolución de puzzles presuponen de forma explícita o implícita que todas las piezas del conjunto inicial pertenecen al objeto, y que éste se reconstruye de forma completa. En estas condiciones, y si el número de fragmentos es bajo, la reconstrucción resulta muy eficiente. Sin embargo, en el escenario y con el tipo de restos que se presentan en este trabajo, ninguna de estas condiciones se cumple. Muchos de los fragmentos, se han perdido o no pueden utilizarse debido a su tamaño o estado de conservación. Por ello, en la mayoría de los casos no es posible completar la reconstrucción. Además, el número de piezas rescatadas es muy elevado: en una base de datos se han almacenados cientos de miles de fotografías de fragmentos recogidos de diferentes zonas del yacimiento. Piezas que serán pertenecientes a diferentes composiciones.

El número de imágenes a procesar es quizá el aspecto más crítico a la hora de conseguir una técnica eficiente de reconstrucción. La comparación por pares de la fractura y el contenido de los fragmentos es computacionalmente muy costosa. Con la tecnología de un ordenador actual, se necesitan minutos para procesar y comparar tan sólo dos fragmentos. Si este proceso se repitiese con todas las piezas, el problema crecería de forma factorial y el plazo de resolución no sería razonable. Resulta obligado reducir el conjunto inicial de fragmentos candidatos a la reconstrucción empleando otras técnicas.

En el presente trabajo, proponemos una versión modificada de conocidas técnicas de correspondencia de imágenes (CBIR – *Content-Based Image Retrieval*). Este algoritmo actúa como filtro para la preselección de fragmentos en base a su similitud visual. La forma de la fractura y el color de las piezas se ignoran debido a que la distinta conservación de los fragmentos hallados puede hacer variar estas características.

El documento se ha organizado como sigue. En la sección 2, resumimos los aspectos principales de las tecnologías CBIR actuales. Las propiedades visuales y ciertos aspectos relacionados con la reconstrucción de atauriques se describen en la sección 3. En las siguientes, se presenta el algoritmo CBIR propuesto y se analizan los resultados obtenidos. Y por último, en la sección 6 resumimos los trabajos actualmente en curso para mejorar las prestaciones de esta técnica.

2. TECNOLOGÍAS DE CORRESPONDENCIA DE IMÁGENES (CBIR)

Las técnicas de correspondencia (CBIR) se diseñaron inicialmente para localizar imágenes parecidas en una base de datos. Se selecciona una imagen: la consulta; y la búsqueda por correspondencia devuelve un conjunto ordenado de imágenes con similitud decreciente. Las propiedades que se tienen en cuenta en la comparación son los colores, las formas y las texturas. Estas tres características están relacionadas de forma indirecta y abstracta con ciertos descriptores matemáticos (Deselaers *et al.*, 2004) como los histogramas de tinta (que representan la distribución o la frecuencia de aparición de los colores presentes en la imagen), los descriptores de Fourier (relacionados con la periodicidad y orientación de los elementos interiores de la escena), los coeficientes Haar (que codifican la distribución de gradientes de luces en diversas escalas), el canal H (la tinta en el espacio de color HLS) entre otros.

Para que la búsqueda y comparación de imágenes sea eficiente y se ejecute de forma rápida, las técnicas de correspondencia separan la fase de procesado y análisis de la de comparación de las imágenes. Al catalogar cada imagen, ésta se procesa y analiza, y se salva en la base de datos junto con un conjunto de descriptores numéricos. En esta fase, que se ejecuta una sola vez por cada imagen de la base de datos, se aplican los algoritmos de mayor carga computacional. Cuando se ejecuta una búsqueda, la imagen de consulta se compara con las de la base de datos empleando únicamente los descriptores y no las imágenes. Gracias a ello, se reduce el tiempo de respuesta en varios órdenes de magnitud (de minutos a milisegundos en cada comparación).

Sin embargo, la similitud de dos imágenes en el sentido humano está relacionada con la información semántica del contenido de la imagen, es decir, con los elementos que el ojo humano identifica en la escena. Muchos investigadores han señalado que los descriptores empleados en las técnicas de correspondencia no contienen, o al menos no codifican, el contenido semántico de las imágenes (Enser, 2003; Traina *et al.*, 2006). Por ejemplo, en la lámina I se muestran dos escenas con pequeñas diferencias (a) y (b) y otra completamente distinta (c). Un sistema de correspondencia basado en los descriptores de Haar señalaría la escena (a) como más parecida a la (c) que a la (b).

Muchos errores pueden evitarse controlando la iluminación y el fondo de la escena, otros fallos requieren modificaciones que dependen de las características de los objetos, y en algunos casos incluso nuevos algoritmos de procesado. En este trabajo demostramos que a pesar de todo, las técnicas de correspondencia son una tecnología útil para ayudar en la reconstrucción de atauriques y, que personalizando correctamente los algoritmos, existe una alta probabilidad de que el subconjunto seleccionado contenga muchos fragmentos similares al de la consulta.

3. PROPIEDADES VISUALES DE LOS ATAURIQUES

El ataurique es una decoración en yeso con motivos de origen vegetal, muy utilizados en el arte mudéjar. Las imágenes de los atauriques presentes en este trabajo provienen del catálogo de Madinat al Zahra, una antigua ciudad árabe situada en las afueras de

Córdoba (España). De las ruinas de Madinat al Zahra se han catalogado hasta la fecha cientos de miles de fragmentos, aunque esto, tan sólo representa un pequeño porcentaje de los realmente existentes. La lámina II muestra una vista panorámica del yacimiento y un ejemplo del estado de los fragmentos que se pueden encontrar en una de las estancias.

Un aspecto que influye de forma crítica en los resultados es el estado de conservación de los fragmentos. En general, se encuentran muy deteriorados. Las fracturas no son limpias, presentan bordes biselados y erosionados, con zonas de sombra a distintas alturas. Todo ello induce a trazar falsos bordes de fractura que reducen e impiden un correcto alineamiento con otros fragmentos. El moho está presente en muchas piezas. Su presencia introduce sombras y cambios de color que crean falsos contornos y falsas texturas, y por ello, se limpian antes de ser fotografiadas. Tras la limpieza es habitual rotular los fragmentos con un número de identificación.

La mayoría de los algoritmos de procesado localizan falsos contornos en las rotulaciones y en las manchas o restos de moho. Es muy importante eliminar ambos antes de procesar las imágenes. En el trabajo hemos desarrollado un algoritmo de limpieza semi-automático muy simple y rápido que redibuja sobre las etiquetas patrones de pixels estadísticamente similares a los del área adyacente. La lámina III muestra ejemplos de todas estas situaciones.

Del estudio de luces y color de las imágenes de los fragmentos, se deduce que poseen poco contraste y que los colores son muy uniformes. Como consecuencia, resulta casi imposible extraer propiedades locales significativas tales como esquinas, líneas, flancos o vértices. El resultado es el mismo independientemente del algoritmo y el umbral utilizado. Por tanto, la identificación de dibujos u ornamentos a partir de estas propiedades locales resulta extramadamente compleja. La figura 1 ilustra estos resultados.

El algoritmo de correspondencia propuesto en el trabajo, utiliza diversos descriptores tanto locales como globales; buscando que la complementariedad de los descriptores mejore los resultados en las comparaciones.

4. CÁLCULO Y COMPARACIÓN DE DESCRIPTORES (CBIR)

Durante el proceso de captura fotográfica es extremadamente importante controlar las dimensiones de la escena para poder diferenciar de forma adecuada la escala de los fragmentos. El tamaño relativo de esos fragmentos debe estar presente a lo largo de todo el procesado para evitar confusiones en la fase de correspondencia (comparación) final. Por ejemplo, una hoja grande no puede hacerse corresponder con una pequeña, aunque sean el mismo motivo. Este efecto es muy frecuente cuando se emplea el zoom de la cámara. Por tanto, se recomienda emplear ópticas fijas y escenas de tamaño predefinido.

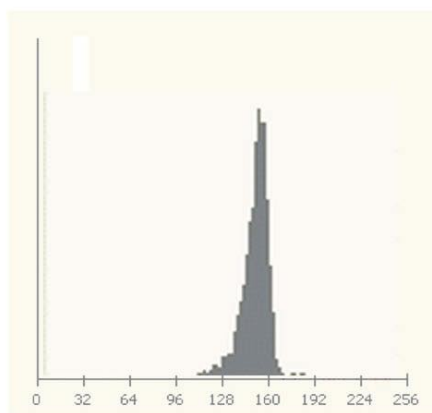
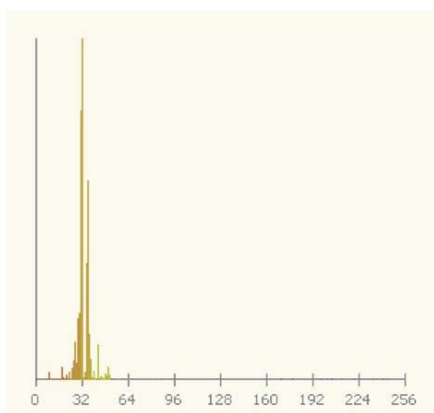
El algoritmo de cálculo de descriptores consta de tres fases: preprocesado de la imagen, texturización y extracción de características. El principal objetivo de los algoritmos de pre-procesado es reducir el tamaño y la resolución de las imágenes y separar el fondo del fragmento. Este proceso incrementa la eficiencia global del algoritmo y reduce los requisitos hardware del ordenador. Las imágenes generadas



Distribución de color



Distribución de luces



Singularidades locales

Fig. 1.—Resultados del análisis de color, luz y puntos singulares.

en la cámara son excesivamente grandes para manejarlas de forma eficiente. Pero además, la mayoría de las técnicas de análisis local de características fallan con imágenes de alta resolución debido a que existe muy poca variación en las propiedades de pixels adyacentes.

4.1. Fase de texturización

La síntesis de textura se aplica en este problema para eliminar o reducir la información de la forma de la fractura se codifica en los descriptores. La texturización se hace de modo que se traslada la información del contenido interior sobre el espacio de la escena no ocupado por el fragmento (fondo), convirtiendo de este modo, la información de contenido en una propiedad global.

El algoritmo utilizado se basa en el propuesto por Paul Francis Harrison (Harrison, 2005). Pixels y texturas se toman del interior del fragmento y se copian sobre las áreas exteriores. Una función de evaluación decide si los pixels a generar se seleccionan en función de sus vecinos inmediatos o de otras partes del fragmento escogidas de forma aleatoria. Para eliminar el efecto que los contornos y biseles provocan sobre la textura final, el área de referencia se reduce despreciando la información cercana a los bordes de la textura. Los pixels descartados son redefinidos como fondo, y por tanto, se incluyen en el proceso de texturización. La lámina IV ilustra el resultado de aplicar esta técnica sobre dos fragmentos.

4.2. Cálculo de los descriptores de Fourier y Haar

Las transformadas de Fourier y Haar son dos operaciones clásicas con propiedades bien conocidas que se aplican normalmente sobre las imágenes originales. En nuestro caso, estas transformaciones, y los descriptores que se extraen de ellas, se aplicarán sobre las imágenes texturizadas. La transformada de Fourier revela propiedades de las texturas como periodicidad y granularidad. Ambas están relacionadas directamente con el tamaño y la repetición del elemento principal de la textura. Estas propiedades se codifican midiendo la energía sobre diferentes áreas de la transformada de Fourier compleja. Como resultado, se obtienen un vector que resume y permite comparar la similitud de estas propiedades entre dos imágenes.

La transformada Haar es una técnica multiresolución capaz de representar patrones no periódicos en diferentes escalas. Los coeficientes Haar más significativos se utilizan para codificar esos patrones. En nuestro caso, gracias a la texturización previa, se reducen los valores de los coeficientes relacionados con el área que ocupa la pieza en la escena, y que contendrían la información de forma y fractura. De este modo, se potencia la información relativa al ornamento interior de la pieza. De nuevo, el resultado es un vector de coeficientes con el que comparar.

4.3. Algoritmos de comparación y búsqueda

Se han definido diferentes métricas para los descriptores de Fourier y Haar. Para el primero se emplea la distancia euclídea de los dos vectores. Los coeficientes de la imagen de referencia se comparan con los todos y cada uno de los descriptores de cada imagen de la base de datos. Como resultado se crea una lista ordenada por similitud.

Para comparar los descriptores de Haar se emplean sólo los coeficientes más significativos. Como medida de similitud entre dos fragmentos se utiliza el número de coeficientes que coinciden entre ambos. De nuevo, tras comparar la imagen de referencia con las de la base de datos, se obtiene una lista ordenada por semejanza.

El resultado final de la comparación de dos fragmentos se obtiene fusionando las dos listas anteriores. En el presente trabajo, proponemos una variante del algoritmo de votación conocido como Recuento Borda (Van Erp, 2000). Asociando un peso de voto a la posición de cada una de las listas anteriores, la lista resultante contiene por orden los resultados más votados.

4.4. Optimización y dependencias de la escala

Como ya se ha apuntado, la dependencia de la escala es un punto crítico en el proceso de comparación, ya que está directamente relacionado con el tiempo de cálculo e influye en la tasa de éxito de los resultados. Dos fragmentos de distinto tamaño, fotografiados con diferente óptica, tendrán similares dimensiones en sus respectivas fotografías, y por tanto los algoritmos de correspondencia los considerarán semejantes. La ratio escena/imagen (en adelante, la escala de la imagen) debería ser constante, o al menos, debe ser conocida. El método más sencillo para lograrlo es emplear siempre las mismas dimensiones de escena, que estará determinada por el tamaño de los fragmentos más grandes que puedan hallarse.

Sin embargo, emplear una escena común tanto para piezas grandes y pequeñas reduce la efectividad de los algoritmos de comparación e incrementa el tiempo de procesado. Por ejemplo, en el caso de fragmentos pequeños, se debe rellenar con textura grandes áreas o espacios vacíos. El algoritmo de texturización es muy pesado computacionalmente, y además, cuanto más pequeño es el patrón de relleno (la pieza) en relación al área total, mayor es la degradación que presenta el resultado. Como consecuencia, grandes áreas a texturizar aumentan el tiempo de cálculo y reducen la calidad de la textura. Para evitar ambos inconvenientes, el algoritmo de cálculo de los descriptores de Fourier se ha modificado para incluir dependencia del tamaño de la escena y múltiples escalas (aunque prefijadas). La posición y focal de la cámara deben ajustarse de tal forma que contengan exactamente las dimensiones de las escenas predefinidas. La cuantización o resolución espacial empleada debe preservar las proporciones de aspecto del sensor de la cámara (4:3 o 2:3). La tabla I muestra la relación entre las escenas permitidas y el tamaño de las imágenes y de las escenas y las resoluciones seleccionadas para cada una.

TABLA 1
TAMAÑOS DE LA ESCENA Y DE LA IMAGEN

<i>Escena</i>	<i>Tamaño de la imagen</i>
120 x 90 cm	256 x 256 px
60 x 45 cm	128 x 128 px
30 x 22,5 cm	64 x 64 px

Con el objetivo de amplificar la información sobre el contenido, la texturización se calcula sobre una imagen de tamaño doble a la inicialmente asignada por las dimensiones de la escena y de modo que el original se sitúe en el centro. De ella se extrae una segunda imagen de textura con las dimensiones originales. Finalmente los descriptores de Fourier se calculan sobre ambas. La figura 2 ilustra esta técnica.

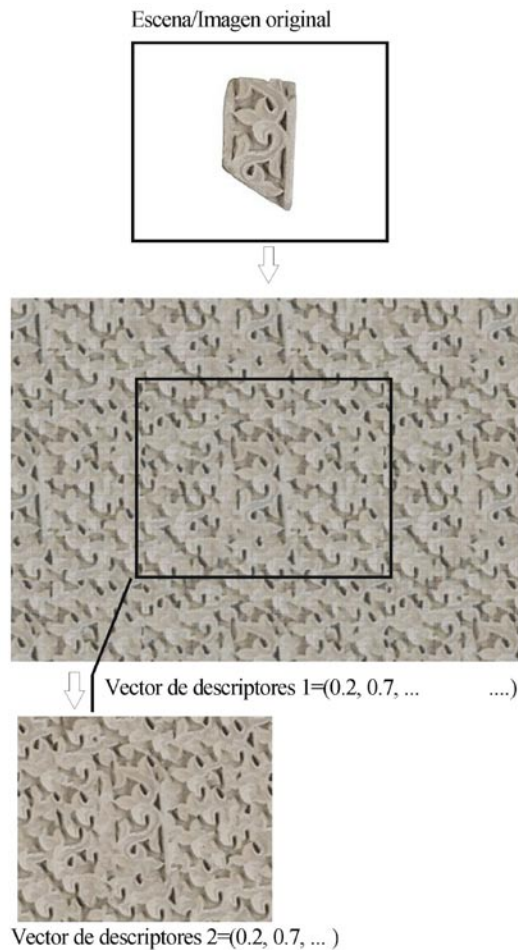


Fig. 2.—Proceso de generación de firmas a diferentes escalas.

La dimensión, o número de coeficientes de los vectores de Fourier, es diferente para cada escena. Cuando se comparan dos fragmentos a través de los descriptores, éstos deberán tener igual dimensión, es decir, deben corresponder a escenas e imágenes de igual tamaño, y por ende, corresponderán a piezas de tamaño similar.

5. RESULTADOS

Para medir las prestaciones de la técnica desarrollada hemos diseñado dos pruebas. La finalidad de la primera es medir su capacidad de agregación, es decir la capacidad de reconocer como similares fragmentos que posean el mismo contenido. Para ello se parte de dos grupos de imágenes procedentes de dos conjuntos arqueológicos diferentes. Un grupo de 120 imágenes y otro de 12. La prueba consistía en seleccionar fragmentos no catalogados del segundo conjunto y ver si el sistema era capaz de clasificar la pieza correctamente, a partir de los resultados de similitud proporcionados por nuestro CBIR. El éxito ha sido superior al 90% en los casos estudiados.

La segunda prueba consiste en medir la capacidad de clasificar fragmentos con distintos motivos. La base de datos de pruebas utilizada consta de 70 imágenes organizadas en 7 grupos. Cada grupo posee 10 imágenes de fragmentos con contenido similar. La lámina V muestra alguna de las piezas de esas cuatro categorías.

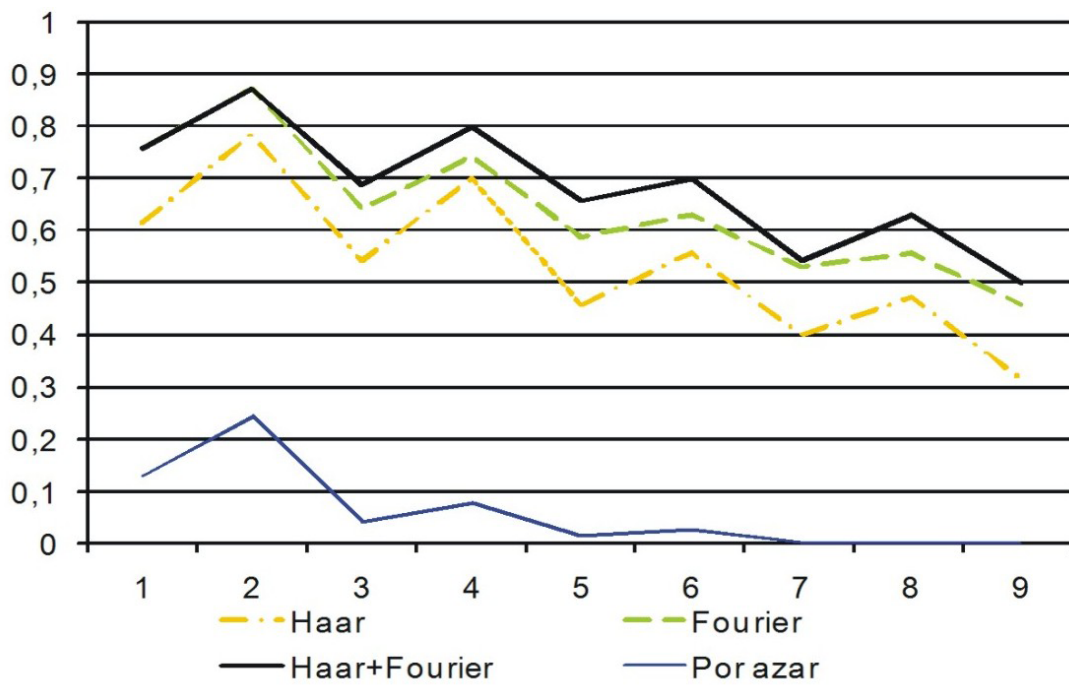


Fig. 3.—Éxitos en calculados.

Los resultados se ofrecen en la figura 3. En ella se muestra el grado de éxito (en tanto por uno) del clasificador en función de considerar los primeros “n” resultados. En dicha figura se observa que determinar la categoría del fragmento alcanza casi un 90% cuando se toma la información de los dos fragmentos más similares de la base de datos. Igualmente se muestra que el resultado combinado Fourier+Haar da un resultado superior que emplear cada método por separado y, en cualquier caso, los resultados son muy superiores a emplear el azar.

6. FUTUROS TRABAJOS

Para aumentar las prestaciones del algoritmo, actualmente estamos testando una técnica de categorización que consiste en comparar la imagen de consulta con un conjunto de imágenes de referencia previamente clasificadas en categorías. Cada categoría contiene fragmentos diferentes pero con similares contenidos ornamentales. De la comparación con éstas puede extraerse información que permite aplicar una heurística que ajuste los pesos de los diferentes algoritmos para mejorar la tasa de aciertos, empleando, por ejemplo, la distribución de frecuencias de las categorías seleccionadas.

También se están estudiando otras técnicas de la comparación de singularidades locales como la transformada SIFT, técnicas de comparación de curvas basadas en trazos y la selección de nubes de puntos imprecisos e incompletos.

AGRADECIMIENTOS

Todos los trabajos desarrollados han sido financiados por ICOSIS (www.icosis.es) una empresa de software situada en Sevilla (España) y la Agencia de Innovación de Andalucía (IDEA). Las imágenes han sido cedidas por la Consejería de Cultura (Junta de Andalucía) y el centro arqueológico de Madinat al-Zahra en Córdoba (España).

BIBLIOGRAFÍA

- CARLINI, A., CONVERSANO, E. y TEDESCHINILALLI L. (2008): “Mathematics and Archaeology”, *7th Int. Conf. APLIMAT 2008*, Bratislava, Slovak
- CARSON, C., THOMAS, M., BELONGIE, S., HELLERSTEIN, J.M. y MALIK, J. (1999): “Blobworld: A system for region-based image indexing and retrieval”, *Int. Conf. Visual Information Systems*, pp. 509–516, Amsterdam, The Netherlands, June 1999.
- DESELAERS, T., KEYSERS, D. y NEY, H. (2004): “Features for Image Retrieval: A Quantitative Comparison”. *Pattern recognition. 26th DAGM Symposium*, Tübingen, Germany, August 30 - September 1, 2004.
- ENSER, P. (2003): “Towards a Comprehensive Survey of the semantic Gap in Visual Image Retrieval”, *Image and Video Retrieval, Lecture Notes in Computer Science*, pp. 163-128.
- FALOUTSOS, C., BARBER, R., FLICKNER, M., HAFNER, J., NIBLACK, W., PETKOVIC, D. y EQUITZ, W. (1994): “Efficient and effective querying by image content”, *Journal of Intelligent Information Systems* 3(3/4):231–262, July 1994.
- KEYSERS, D., GOLLAN, C. y NEY, H. (2004): “Classification of medical images using non-linear distortion

- models”, In *Bildverarbeitung für die Medizin*, pp. 366–370, Berlin, Germany, Mar. 2004.
- REDDY, B.S. y CHATTERJI, B. (1996): “An FFT-based technique for translation, rotation and scale invariant image registration”. *IEEE Trans. Image Proc.*, 5(8):1266–1271, Aug 1996.
- TAMURA, H. , MORI, S. y YAMAWAKI, T. (1978): “Textural features corresponding to visual perception”, *IEEE Transactions on Systems, Man and Cybernetics* 8 (6): 460 – 473, Jun 1978.
- TRAINA, A.J-M., MARQUES, J. y TRAINA, C. Jr. (2006): “Fighting the Semantic Gap on CBIR Systems through New Relevance Feedback Techniques”, *19th IEEE Symposium on Computer-Based Medical Systems (CBM 06)*, pp. 881-886.

© Theo Pavlidis (<http://theopavlidis.com>)



(a)

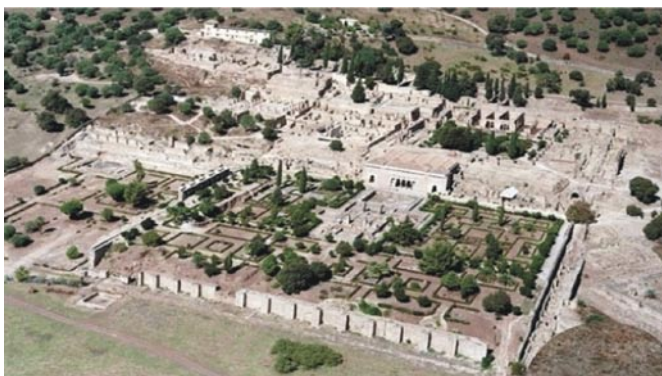


(b)



(c)

Lám. I.—La imagen (b) es una versión más luminosa de la (a), sin embargo se clasifica como más parecida a (c) (© Theo Pavlidis (<http://theopavlidis.com>))



Lám. II.—Vista panorámica de Madinat al Zahra, y de los fragmentos de atauriques. (C) Junta de Andalucía.



Lám. III.—Estado de los fragmentos tras la limpieza, y resultado del borrado digital.



Lám. IV.—Resultados del algoritmo de texturización.



Lám. V.—Ejemplos de categorías utilizadas en las pruebas.